Claims

What is claimed is:

1. A method for use in resource discovery of establishing a semantic correspondence between a first set of labels and a second set of labels, the method comprising the steps of:

obtaining one or more examples and a classifier for the first set of labels, and one or more examples and a classifier for the second set of labels;

using the classifier associated with the first set, trained on examples from the first set, to classify the second set thereby generating classification results for the second set of labels, and using the classifier associated with the second set, trained on examples from the second set, to classify the first set thereby generating classification results for the first set of labels; and

generating label association rules based on the classification results for the first set of labels and the classification results for the second set of labels, a label association rule having a semantic correspondence measure of confidence associated therewith.

2. The method of claim 1, further comprising the step of identifying one or more label association rules, from the label association generated rules, which have a measure of confidence not below a given threshold value.

3. The method of claim 2, wherein the one or more identified label association rules are used in a resource discovery operation associated with a requested search.

4. The method of claim 3, wherein the resource discovery operation is distributed.

5. The method of claim 1, wherein the semantic correspondence measure of confidence for a label from the first set with respect to a label of the second set is a sum

of respective confidence measures associated with classification of the one or more examples associated with the sets of labels.

6. The method of claim 1, wherein the label association rules generation step further comprises the steps of:

generating rules associating labels of the first set to labels of the second set;

generating rules associating labels of the second set to labels of the first set; and

computing the semantic correspondence measure of confidence for a label i from one of the sets and a label J from the other of the sets as the product of a semantic correspondence confidence measure associated with the label i implying the label J and a semantic correspondence confidence measure associated with the label J implying the label i.

7. The method of claim 1, wherein classification of labels is supervised.

8. The method of claim 7, wherein supervised classification is performed in accordance with one of a Bayes classification algorithm, a Perceptron classification algorithm, a k-nearest-neighbor classification algorithm, a linear discriminant function classification algorithm, and a neural networks classification algorithm.

9. Apparatus for use in resource discovery of establishing a semantic correspondence between a first set of labels and a second set of labels, the apparatus comprising:

at least one processor operative to: (i) obtain one or more examples and a classifier for the first set of labels, and one or more examples and a classifier for the second set of labels; (ii) use the classifier associated with the first set, trained on examples from the first set, to classify the second set thereby generating classification results for the second set of labels, and use the classifier associated with the second set,

trained on examples from the second set, to classify the first set thereby generating classification results for the first set of labels; and (iii) generate label association rules based on the classification results for the first set of labels and the classification results for the second set of labels, a label association rule having a semantic correspondence measure of confidence associated therewith.

10. The apparatus of claim 9, wherein the at least one processor is further operative to identify one or more label association rules, from the label association generated rules, which have a measure of confidence not below a given threshold value.

11. The apparatus of claim 10, wherein the one or more identified label association rules are used in a resource discovery operation associated with a requested search.

12. The apparatus of claim 11, wherein the resource discovery operation is distributed.

13. The apparatus of claim 9, wherein the semantic correspondence measure of confidence for a label from the first set with respect to a label of the second set is a sum of respective confidence measures associated with classification of the one or more examples associated with the sets of labels.

14. The apparatus of claim 9, wherein the label association rules generation operation further comprises: (i) generating rules associating labels of the first set to labels of the second set; (ii) generating rules associating labels of the second set to labels of the first set; and (iii) computing the semantic correspondence measure of confidence for a label i from one of the sets and a label J from the other of the sets as the product of a semantic correspondence confidence measure associated with the label i implying the

label J and a semantic correspondence confidence measure associated with the label J implying the label i.

15. The apparatus of claim 9, wherein classification of labels is supervised.

16. The apparatus of claim 15, wherein supervised classification is performed in accordance with one of a Bayes classification algorithm, a Perceptron classification algorithm, a k-nearest-neighbor classification algorithm, a linear discriminant function classification algorithm, and a neural networks classification algorithm.

17. An article of manufacture for use in resource discovery of establishing a semantic correspondence between a first set of labels and a second set of labels, comprising a machine readable medium containing one or more programs which when executed implement the steps of:

obtaining one or more examples and a classifier for the first set of labels, and one or more examples and a classifier for the second set of labels;

using the classifier associated with the first set, trained on examples from the first set, to classify the second set thereby generating classification results for the second set of labels, and using the classifier associated with the second set, trained on examples from the second set, to classify the first set thereby generating classification results for the first set of labels; and

generating label association rules based on the classification results for the first set of labels and the classification results for the second set of labels, a label association rule having a semantic correspondence measure of confidence associated therewith.

18. In a client-server arrangement, a method for use in resource discovery of establishing a semantic correspondence between a first set of labels associated with a query provided at a client device and at least a second set of labels associated with at least

one information repository associated with at least one server, the method comprising, at one of the client device and the server, the steps of:

obtaining one or more examples and a classifier for the first set of labels associated with the client query, and one or more examples and a classifier for the second set of labels associated with the information repository;

using the classifier associated with the first set, trained on examples from the first set, to classify the second set thereby generating classification results for the second set of labels, and using the classifier associated with the second set, trained on examples from the second set, to classify the first set thereby generating classification results for the first set of labels; and

generating label association rules based on the classification results for the first set of labels and the classification results for the second set of labels, a label association rule having a semantic correspondence measure of confidence associated therewith.